Analysis of fast coding unit size decision algorithms for HEVC

Dr. PL. Chithra∗1, A. Roselin Clara2
1 Associate Professor, University of Madras
2 Research Scholar, University of Madras
1 chitrasp2001@yahoo.com 2 roselinclara.research@gmail.com

March 21, 2018

Abstract
HEVC (High Efficiency video coding) standard produces better compression rate compared with its previous standards H.264/AVC (Advanced Video Coding). HEVC employs hetero block sizes and its coding structure has three main units called coding unit (CU), prediction unit (PU), and transform unit (TU). Coding unit has to be split into multiple coding block and prediction type (intra or inter prediction with 35 intra prediction modes) is decided at each coding block. Splitting of CU using quad tree structure helps in high compression efficiency ratio but involves lot of computational complexity as all combinations of mode candidates are calculated to get minimum RD (rate distortion) cost. The computational complexity on encoder were due to the optimization processing in the efficient coding tools, especially the rate distortion (RD) optimization on coding unit (CU), prediction unit, and transform unit. To reduce the computational complexity and achieve high coding efficiency many fast CU size decision algorithms for intra and inter coding which reduce the depth level were proposed. This paper reviews those algorithms and analysis its performance.

Key Words: HEVC, coding unit, quad tree structure, mode decision, machine learning.


1 INTRODUCTION

Today’s world is full of digital applications and they are of high definition. To store the images and videos with high definition needs lots of memory space and to transmit to other systems need lots of bandwidth. To reduce the space and bandwidth we use compression techniques. Lots of compression encoders for image and video are available in the market. The latest encoder for video is HEVC which is mainly for high definition. HEVC achieves high coding efficiency compared to its previous standard H.264/AVC. The coding efficiency is the bitrate necessary to represent the video content with good video quality and it’s efficient in HEVC.

High Efficiency Video Coding (HEVC) is a new video compression standard developed jointly by ITU-T Video Coding Experts Group (VCEG) and ISO/IEC Moving Pictures Expert Group (MPEG) through their Joint Collaborative Team on Video Coding (JCT-VC). Fig 1 shows the HEVC hybrid encoder. HEVC employs a quad tree structure in which the smallest coding unit (CU) can be a minimum of 8 × 8 pixels and the largest coding unit (LCU) can be a maximum of 64 × 64 pixels. The tree structured CU completely breaks the normal procedure of 16 × 16 macroblock (MB) coding architecture found in H.264/AVC[1]. This paper focuses on different algorithms to decide on the depth level of CU based on minimum RDO cost in HEVC. The rest of the paper is organised as follows: Section II describes about coding, prediction, transform unit Section III presents a brief literature review on CU size decision algorithms. Experimental results are analysed in Section IV to demonstrate the efficiency of the algorithms, while Section V concludes the work.
2 Related Work

Coding and Prediction unit

The main advantage of HEVC over its previous standard H.264/AVC is that it has quad tree structure block partitioning instead of macroblock. Video are divided into slices. Then the slices are partitioned into many coding tree units (CTU). CTU is the largest block in HEVC and can be of size 64 X 64, 32 X 32, 16 X 16 with depth as 0. A CTU can be split into four sub blocks with depth one larger than the previous level. Each sub block can be recursively split to form a coding tree structure until the desired depth is reached. This paper analysis the different algorithms available to decide the depth level where the rate distortion (RD) cost is minimum. Each leaf node of the coding tree structure is called as CU (coding unit) and can be split into prediction unit (PU). Fig 2 shows the coding tree structure of HEVC. [7]
Each prediction unit has its own coding information and can have either intra or inter prediction. Intra prediction uses 35 intraprediction modes and have many algorithms in the literature to decide the minimum number of prediction modes based on the RD cost and rough mode decision (RMD) to choose fewer candidates out of 35 intra prediction modes. Inter prediction coding uses motion vector. The coding tree structure and multiple PU modes allow better flexibility for block partitioning and highly increase the coding efficiency. Fig 3 shows the different partition for PU [7].

The residual signal of the intra- or inter-picture prediction, which is the difference between the original block and its prediction, is transformed by a linear spatial transform in the TU (Transform unit). The transform coefficients are then scaled, quantized, entropy coded, and transmitted together with the prediction information.

Recently, many researchers devoted their efforts on the reduction of the computational complexity for HEVC intra prediction, and the released algorithms could be roughly classified into two main categories, which are fast intra prediction mode decision and fast CU/PU size decision, respectively. The main idea of the former is to reduce the candidate modes for rough mode decision (RMD) method and most probable modes (MPM) by commonly utilizing the direction information of image edge and the correlation of the modes. This kind of algorithms has positive effects on accelerating the intra prediction process, but they also need to search all CU depths to find the optimal quad-tree partition. For the latter, some researchers proposed fast CU size decision algorithms based on the
analysis of video information to reduce the computational complexity of a typical HEVC encoder radically.[5] The following section review the different algorithms in the literature.

3 Review on CU size decision algorithms

3.1 Coding unit size decision with intra coding

A fast CU size decision algorithm for HEVC intra coding was proposed to speed up the process by reducing the number of candidate CU sizes required to be checked for each tree block. The novelty of the proposed algorithm lies in the following two aspects: an early determination of CU size decision with adaptive thresholds is developed based on the texture homogeneity and a novel bypass strategy for intra prediction on large CU size was proposed based on the combination of texture property and coding information from neighbouring coded CUs. Experimental results show that the proposed effective CU size decision algorithm achieves a computational complexity reduction up to 67%, while incurring only 0.06-dB loss on peak signal-to-noise ratio or 1.08% increase on bit rate compared with that of the original coding in HM. [1]

3.2 Coding unit for intra coding with texture complexity analysis

It is obvious that the homogeneous region is more inclined to be coded by large block and the region with detailed texture is usually coded by small block in image encoding process. Enlightened from this phenomenon, the texture complexity of coding block is applied to determine whether it should be split or skipped in each depth. The most important issue of CU size selection process is the definitions of unsplit depth and skip depth. With abundant experimental research, they found that the thresholds of unsplit depth and skip depth should be made adaptive to the quantization parameter (QP) and coding depth. It indicates that the unsplit depth is the exponential function of the depth and the nonlinear function of the QP.
The performance of the entire proposed algorithm proves that the proposed algorithm shows its advantages both in coding efficient and complexity. About 56% on average (up to 68.2%) computational complexity is saved with just 1.0% BD-Rate Y increase which could be negligible. Moreover, the coding performances for the chrominance component are maintaining very well (only 0.1% and 0.2% increase for BD-Rate U and BD-Rate V, respectively), since the proposed method just slightly influence the CU splitting and Intra Derived mode decision for Intra chrominance coding. [2]

3.3 Coding unit selection method for inter coding based on pyramid motion divergence

In this paper, a pyramid motion divergence (PMD) based method is proposed to early skip the specific inter CUs in HEVC. PMD feature is derived to be used for CU selection. For each CU, the PMD is evaluated as the variance of the optical flows of the current CU and the sub-CUs. The implicit relation between the motion divergence and the R-D cost is investigated. A nearest neighbouring (NN) algorithm is used to determine the splitting type of each CU to reduce the computational complexity of RDO. The PMD feature of a CU is calculated as the variance of the pixel MVs in and the variance of pixel MVs in the four sub-CUs. Thus, PMD can be denoted as,

\[ P_x = (V_{ar,M_X}, V_{ar,M_0}, V_{ar,M_{x1}}, V_{ar,M_{x2}}, V_{ar,M_{x3}}), \]

where,

\[ V_{ar,M_X} = \sum_{x \in X} \| m - m_x \|^2 \text{ and } X_i \text{ is the } i^{th} \text{ sub-CU of } X \]  (1)

The proposed system search the similar CUs to select the optimum CU splitting. The efficiency of the models is evaluated with respect to the time savings and the coding quality. The experimental results show that the proposed algorithm reduces the encoding time significantly and also has a negligible coding loss. [3]

3.4 Coding unit size decision with machine Learning

The paper propose a fast algorithm for the CU partitioning process of the HEVC encoder using machine learning methods. A com-
plexity measure based on the Sobel operator and rate-distortion costs are defined as features for the algorithm. A CU size can be determined early by employing Fishers linear discriminant analysis and the k-nearest neighbours classifier. The statistical data used for the proposed algorithm is updated by adaptive online learning phase. The experimental results show that the proposed algorithm can reduce encoding time by approximately 54.0\% with a 0.68\% Bjntegaard-Delta bit-rate increase. [4]

3.5 Coding unit decision with complexity classification using machine learning (CCML)

This paper proposes an adaptive fast CU size decision algorithm, named as CCML (Complexity Classification using Machine Learning), for HEVC Intra frame coding. Initially, the quad tree-based CU size decision process and the relationship between CU partitions and image features are analysed. And then, a three-output classification model is constructed based on the CU complexity by using SVM technique, which guarantees a relatively higher classification accuracy than a simple binary classifier. For this proposed algorithm, the offline training mode was chosen to avoid the additional computational burden. Firstly, the training data is generated by encoding multiple video sequences with the original HEVC reference encoder (HM15.0). And then, the offline dual-SVM model is trained with different CARs. Finally, the classification model is loaded into the HEVC reference encoder (HM15.0) to assist in deciding the optimal CU size in advance without performing the full RDO process.

The complexities of each CU was computed and put into the trained SVM models, and it gives the guidance for the CU size decision. The experimental results show that the proposed CCML algorithm succeeds to reduce around 59.66\% encoding time on average compared with the original test model HM15.0, while efficiently maintaining the similar coding performance. [5]
3.6 Coding unit decision using SVM model of machine learning

In this work, a CU splitting early termination algorithm in HEVC was formulized as a binary classification problem on which a support vector machine (SVM) was applied and solved by support vector classification. In order to maintain the RD performance of CU splitting early termination algorithm, RD loss due to misclassification was introduced as weighting factor of training samples in the offline training procedure, with which the training method pays special attention to CUs which are prone to degrade RD performance when using a suboptimal partition. Furthermore, diverse features are considered such as the correlation between CUs both in spatial and temporal domains, prediction errors, motion activities, and RD cost of modes. To select the optimal feature subset, a wrapper feature selection approach was carried out. It embeds the model training into the selection process and simple greedy search was performed based on F-score ranking. Demonstrated by the experimental results, the proposed algorithm can achieve 44.7% reduction in computational complexity with 1.35% BD-Rate increase in Random Access, main configuration and 41.9% complexity reduction with 1.66% BD-Rate increase in Low Delay, main configuration. [6]

4 Experimental Analysis

The performance of the proposed fast mode decision algorithm are analysed with the experimental results by implementing the algorithms in HM Reference Software with the simulation environments. BD-rate and encoding time are employed to measure the coding efficiency and computational complexity. BD-rate and Time, represent the average bitrate differences and encoder time reduction, respectively. Table 1 shows the Comparative analysis of different algorithms reviewed in the paper. CAR in the table refers to Classification accuracy rate and the Time value is negative which is calculated using the formula:

\[
\Delta Time = \frac{Time_{prop} - Time_{HM}}{Time_{HM}} \times 100\% \tag{2}
\]
Table 1: Analysis of the results of the algorithms compared to HEVC reference software (HM)

<table>
<thead>
<tr>
<th>Class</th>
<th>Sequence</th>
<th>Number of Passes</th>
<th>Reduction in Encoding Time</th>
<th>Reduction in BD-Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Foreman/Foreman</td>
<td>1.7</td>
<td>92.99</td>
<td>0.76</td>
</tr>
<tr>
<td>B</td>
<td>Foreman/Foreman</td>
<td>5.5</td>
<td>87.68</td>
<td>0.16</td>
</tr>
<tr>
<td>C</td>
<td>Foreman/Foreman</td>
<td>1.4</td>
<td>94.65</td>
<td>0.87</td>
</tr>
<tr>
<td>D</td>
<td>Foreman/Foreman</td>
<td>1.4</td>
<td>92.86</td>
<td>0.32</td>
</tr>
<tr>
<td>E</td>
<td>Foreman/Foreman</td>
<td>1.4</td>
<td>92.86</td>
<td>0.32</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>1.96</td>
<td>92.52</td>
<td>1.33</td>
</tr>
</tbody>
</table>

The proposed CU selection algorithm can effectively reduce the computational complexity significantly for all sequences, which provides 45 to 60% encoding time reduction in average with negligible increase in BD-Rate. CCML algorithm could also achieve better coding performance compared with the other existing fast CU size decision algorithms in the literature.

5 Conclusion

This paper reviews and analyses different CU size decision algorithms in HEVC for both intra and inter coding. HEVC has high coding efficiency but the computational complexity to decide the proper CU size and the prediction modes for predicting the split CU block is high. Hence many algorithms were proposed to choose the minimum prediction modes with minimum RD cost and decide the depth level of CU split. In future an algorithm for CU split size decision using machine learning will be proposed which reduces the computational complexity and has high coding efficiency.

References


[2] Xingang Liu, Yinbo Liu, Peicheng Wang, Chin-Feng Lai, and Han-Chieh Chao An Adaptive Mode Decision Algorithm Based on Video Texture Characteristics for HEVC Intra Prediction,
IEEE Transactions on Circuits and Systems for Video Technology DOI 10.1109/TCSVT.2016.2556278


